

The Impact of Displayed Awards on the Credibility and Retention of Web Site Information

John Shon MD, Jonathan Marshall, Mark A. Musen MD, PhD
Stanford Medical Informatics, Stanford University School of Medicine,
Stanford, CA 94305-5479

ABSTRACT

Ratings systems and awards for medical Web sites have proliferated, but the validity and utility of the systems has not been well established. This study examined the effect of awards on the perceived credibility and retention of health information on a Web page. We recruited study participants from Internet newsgroups and presented them with information on the claimed health benefits of shark cartilage. Participants were randomized to receive health information with and without a medical award present on the page. We subsequently asked them to evaluate the credibility of the Web page and posed multiple-choice questions regarding the content of the pages. 137 completed responses were included for analysis. Our results show that the presentation of awards has no significant effect on the credibility or retention of health information on a Web page. Significantly, the highly educated participants in our study found inaccurate and misleading information on shark cartilage to be slightly believable.

INTRODUCTION AND BACKGROUND

In recent years, patients have become more active in making decisions about their health care. Increasingly, patients seek information on the Internet to help them with these decisions, and thousands of sites now try to deliver medical information to them¹. Despite the large number of health sites on the Web, it can still be challenging to find sites with high-quality, trustworthy information. Several formal studies have clearly demonstrated problems with the accuracy of medical information on the Internet^{2,3,4,5}. Furthermore, patients often lack the medical knowledge to assess accurately the quality and appropriateness of medical content in Web pages.

Controlling Quality

Several approaches to assist consumers in locating quality sites have been tried, including voluntary publishing standards, such as the Health on the Net (HON) Code of Conduct⁶, trusted source compilations, such as MEDLINEplus⁷, and

“untrustworthy” lists, such as QuackWatch⁸. Perhaps the most common approach is the use of external rating systems. By “external” we refer to the bestowing of awards by an entity not related to the entity producing the Web site. Unfortunately, most of these ratings systems do not have established validity or reliability. In one review, Jadad reported that out of 47 ratings instruments, only 2 used basic publishing criteria as a basis for rating Web sites⁹.

The Use of Ratings Instruments and Awards

Even if a rating instrument is shown to be reliable and valid, authentication of the award is difficult given the nature of the Internet. For example, a site might display an external award or rating without having earned or warranted it. Credible health sources thus may be hindered in allotting awards because they do not wish to police them. Even with strong authentication, it may be difficult or too time-consuming for consumers to determine the significance of an award.

Despite these shortcomings, ratings instruments and awards have proliferated and appear prominently on many consumer health Web sites. Clearly, there is some motivation for placing these awards on the sites. It is very possible that awards are displayed to enhance the credibility of health information on a Web site. Once information is deemed credible, it is likely that consumers will learn, absorb and utilize the information in some manner.

Unfortunately, little previous work has been published on the effects of ratings or awards on consumer health Web sites. In an e-commerce trust study, Cheskin reported that Web-based seals of approval, such as VeriSign, when recognized, do communicate trustworthiness¹⁰. In the Web Credibility Project at Stanford, many factors were found to affect the credibility of Web sites. In this Web-based survey, subjects stated that awards on a Web site increased the believability of that Web site¹¹. Unfortunately, as this was a survey, the information obtained was declarative and not an observed finding. We thus performed a Web-based simulation study to test the hypotheses that awards

can increase the credibility of a Web site and increase the amount people learn from a Web site.

METHODS

Participant Recruitment

Participants were recruited using postings to health-related Internet newsgroups. Potential participants were told that the study would examine how people learn and use medical information on the Internet. Participation was voluntary and did not involve any compensation. The study required low bandwidth for access and was designed to be compatible with most browsers. The protocol underwent Human Subjects approval and consent was obtained as participants entered the study.

Study Content Selection

We chose to use information adapted from a live Web site regarding the medical use of shark cartilage. The use of shark cartilage is often considered to be part of "alternative therapy" or "alternative medicine". Although alternative treatments have been quite popular, their use has been controversial in traditional Western medicine¹². In addition, the passage chosen included a known inaccuracy and misrepresentation: shark cartilage is known to have side effects, including severe gastrointestinal toxicity, and it has been reported to have induced hepatitis^{13,14}. We thus felt that the response to this topic would have high initial variability. We chose to use information regarding frogs as a distracting page, since it also dealt with animals and science. Using a standard graphics program, an award was constructed that connoted medical approval from a conjured physician organization "SLA". Convenience sampling found the logo to have face validity; the logo was deemed credible and was chosen to be more credible than other logos that had been created for the study.

Study Flow

We presented participants with disclaimers and brief instructions before they entered the study. We subsequently presented them with a page containing information on shark cartilage (Figure 1). Subjects were randomized to receive the "SLA" logo in the upper right corner (Figure 1). The layout was designed to remain the same regardless of the presence of the logo. We then presented a page with distracting information about the biology of frogs (Figure 2). Finally, we presented a page that contained questions regarding the credibility and content of the pages and participant demographics (Table 2, Table 3). Answers were obtained using

scales with descriptive statements at the extremes (Table 2).

Figure 1

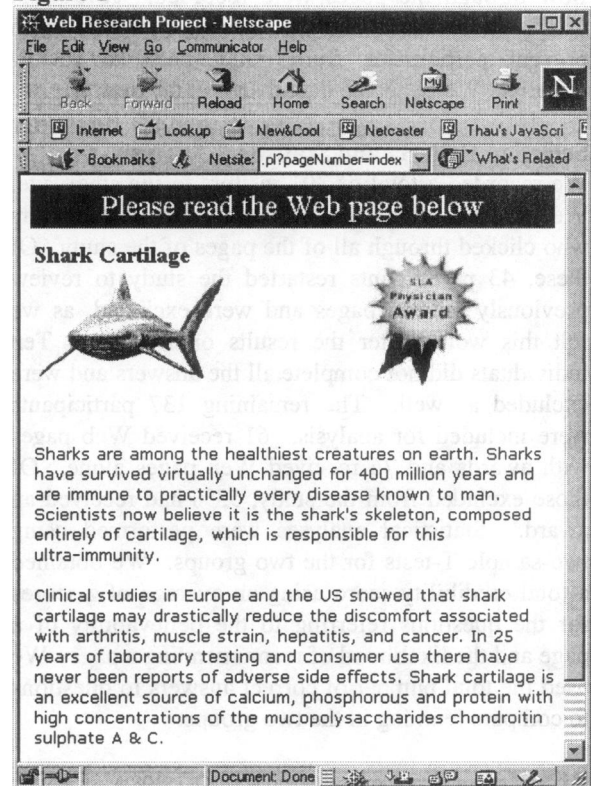
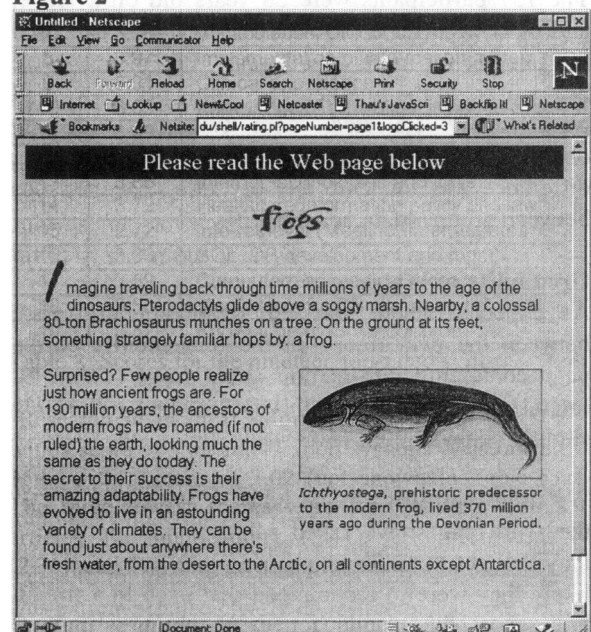


Figure 2



Study Implementation

The study was implemented using Perl and a standard Web server. The time to complete the study and the flow through the pages were recorded. JavaScript was used to close and open browser windows to prevent participants from going back to review material. We collected data over a 7 day period.

Selection Criteria and Statistical Analysis

We recorded a total of 391 visitors to the study site. We considered for inclusion only the 190 visitors who clicked through all of the pages of the study. Of these, 43 participants restarted the study to review previously viewed pages and were excluded, as we felt this would alter the results on learning. Ten individuals did not complete all the answers and were excluded as well. The remaining 137 participants were included for analysis: 61 received Web pages with awards and 76 received Web pages alone. Of those excluded from the study, 53% had received an award. Statistical analyses were performed using two-sample T-tests for the two groups. We obtained a total credibility score using an average of z-scores for the questions referring to the believability of a page and the likelihood of recommending a page. We used the total number of correct answers to questions to compare learning in the two groups.

RESULTS

Participant Characteristics

The 137 participants were 28 years old on average, and 61% were male (Table 1). Over 80% had used the Internet for more than 2 years. Participants were heavy users of the Internet (16.7 hours per week), and were highly educated, with two thirds having obtained a college degree (Table 1). No significant difference was found in any of these characteristics between groups (data not shown).

Credibility of Web pages

As expected, there was no significant difference between the two groups in the believability of the page containing information about frogs (Table 2). Interestingly, they found this information to be only slightly believable.

We used two questions to address the credibility of the shark pages: we asked subjects how believable they thought the information was on a page and how likely they were to recommend that page to a friend. We report the results for each measure, but we combined them into a total credibility score for comparison (Table 2). Both groups found the information on shark cartilage to be slightly

believable, and both groups were neutral to recommending the Web page to a friend. We found no significant difference between the two groups in the total credibility score.

Table 1: Participant Characteristics

Average Age	28
Sex	(%)
Male	61
Female	39
Number of hours/week using the Internet	16.7 hours
Average number of years using the Web	(%)
0-1	9.6
1-2	6.7
2-3	15.6
3-4	15.6
4-5	13.3
5+	39.3
Level of Education	(%)
Some high school	2.9
High School	6.6
Some College	23.3
College Graduate	46.0
Graduate School	21.2

Table 2: Credibility of Web pages

Question	Award Present	No Award Present
	Mean Response	Mean Response
How believable was the information on the shark cartilage Web page? 1 – Not at all believable 7 – Very believable	4.51 (SE = .24)	4.87 (SE = .17)
How believable was the information on the Frogs Web page? 1 – Not at all believable 7 – Very believable	4.51 (SE = .23)	4.89 (SE = .18)
How likely would you be to recommend the shark cartilage page to a person concerned with joint pain? 1 – Would definitely recommend 5 – Would definitely not recommend	2.57 (SE = .15)	2.55 (SE = .13)
T-test	-0.61 (P=0.54)	

Table 3: Accuracy of Responses

Multiple Choice Question	Award Present	No Award Present
	% Correct	% Correct
According to the shark cartilage article, sharks have survived unchanged for how many years?	.49	.33
According to the article, shark cartilage can reduce the discomfort of all the following conditions, EXCEPT:	.39	.47
According to the article, shark cartilage is an excellent source of:	.56	.49
Total % Correct	.48	.43
Total Number Correct	1.44 (SE = .12)	1.29 (SE = .11)
T-test	0.91 (P=0.37)	
Average time spent reading the shark page (seconds)	45	47

Accuracy of Responses

In general, subjects provided correct answers less than 50 percent of the time. We found no significant difference in the two groups in total number of correct answers given, although there was a trend toward more correct answers in the group shown the award. Subjects spent about the same amount of time reading the pages, regardless of whether an award was present.

DISCUSSION

In this study, we found no significant effect of ratings on either the credibility or retention of information from our chosen health Web page. This is somewhat contradictory to the findings of a previous survey done on Web credibility factors. There are many possible explanations that may account for this discrepancy. The previous study was a survey, so although participants stated that awards mattered, in practice they may not. In addition, no specific award was mentioned in the previous study. Cheskin found that the most trusted brands were the most well known ones. As our award and medical organization were fictional, they were probably not recognized. In addition, the study group contained heavy Internet users who may be less likely to believe spurious

awards. A lack of recognition of the award, and perhaps its amateur nature, may have thus decreased any positive effect on the credibility of the shark cartilage Web page.

We also do not know the underlying participant attitudes and knowledge of shark cartilage as a topic. As subjects were recruited from health-related newsgroups, they may have been more familiar with this particular topic, which would tend to decrease any effect in learning. Nonetheless, they answered the three questions with poor accuracy despite a reasonable time spent reading the Web pages.

Believability of Inaccurate Medical Information

Significantly, participants felt that the information presented about shark cartilage was slightly believable. There were clear medical misrepresentations and factual errors in the Web page, yet the information was rated at roughly the same level as the objective frog information. Although this highly educated, heavy Internet-using group was not influenced by the presence of an award, it was fairly accepting of the alternative health information. It is clear that a large percentage of this study group was not able to recognize these inaccurate statements. This is worrisome, for there are potentially significant and detrimental health consequences if actions are taken as a result of reading inaccurate information on the Internet.

Study Limitations

The primary limitations of the study are related to the study design. This was a simulated study, so it is difficult to extrapolate the results to actual behavior on the Web. Participants may not approach the study as seriously as they might when faced with a true medical condition. In addition, the inability to view previously viewed pages is an artificial restriction of the study. As discussed previously, important biases may have been introduced in the population sample, the content, and the award chosen. In particular, the study sample was not optimal. There was a very high "drop-out" rate, although it is unlikely that any selection bias was introduced since roughly half of those excluded were randomized to receive awards. Finally, the study sample consisted of highly educated, heavy Internet users that may not be representative of the general population that seeks medical information on the Internet.

CONCLUSIONS AND FUTURE WORK

Our findings did not show an effect of awards on the credibility or retention of information on a Web page.

By design, the study involved a very limited interaction with information. It ignored and omitted other contextual cues, such as basic publishing criteria, quality of advertisements, quality of design, and referring links, which may be used to judge the credibility of a Web site¹⁰. Many factors contribute to establishing the credibility of a Web site, and the relative contribution of an award relative to these other factors is unknown. In addition, factors other than credibility, such as the relevancy and perceived quality of medical information, also influence the use of information by consumers¹⁵.

We suggest further studies to

- Determine whether well-known awards or logos can affect credibility.
- Determine the importance of other factors that might affect credibility and the relative importance of awards.
- Obtain qualitative information regarding the perception of awards and credibility of information presented.

In conclusion, it is almost assured that the use of ratings systems and awards will continue into the future. By gaining insight into how these awards affect perceptions of credibility, and ultimately the behavior of patients, it should be possible to use awards to the benefit of consumer health. Conversely, the misuse of awards and the publishing of inaccurate information on the Internet will also likely persist, requiring constant vigilance on the part of patients and providers.

Acknowledgements

This work was supported in part by the National Library of Medicine Training Grant NLM LM 07033 (for JS).

References

1. Ferguson T, Kelly WJ. The Ferguson Report; 1999 Jan/Feb. Report No.: Vol 1, No. 1.
2. Impicciatore P, Pandolfini C, Casella N, Bonati M. Reliability of health information for the public on the World Wide Web: systematic survey of advice on managing fever in children at home [see comments]. *Bmj* 1997;314(7098):1875-9.
3. McClung HJ MR. The Internet as a Source for Current Patient Information. *Pediatrics* 1998;101(6)(Jun 1):E2.

4. Biermann JS, Golladay GJ, Greenfield ML, Baker LH. Evaluation of cancer information on the Internet [see comments]. *Cancer* 1999;86(3):381-90.
5. Hersh WR, Gorman PN, Sacherek LS. Applicability and quality of information for answering clinical questions on the Web [letter]. *Jama* 1998;280(15):1307-8.
6. Boyer C, Selby M, Scherrer JR, Appel RD. The Health On the Net Code of Conduct for medical and health Websites [In Process Citation]. *Comput Biol Med* 1998;28(5):603-10.
7. Miller N, Lacroix EM, Backus JE. MEDLINEplus: building and maintaining the National Library of Medicine's consumer health Web service. *Bull Med Libr Assoc* 2000;88(1):11-7.
8. Barrett S. <http://www.quackwatch.com>
9. Jadad AR, Gagliardi A. Rating health information on the Internet: navigating to knowledge or to Babel? *Jama* 1998;279(8):611-4.
10. Cheskin Research. e-Commerce Trust Study. <http://www.studioarchetype.com/cheskin/>
11. BJ Fogg JM, Othman Laraki, Alex Osipovich, Chris Varma, Nicholas Fang, , Jyoti Paul AR, John Shon, Preeti Swani, Marissa Trein. Elements that Affect Web Credibility: Early Results from a Self-Report Study. In: *Human Computer Interaction; 2000; The Hague, The Netherlands; 2000.*
12. Burstein HJ, Gelber S, Guadagnoli E, Weeks JC. Use of alternative medicine by women with early-stage breast cancer [see comments]. *N Engl J Med* 1999;340(22):1733-9.
13. Ashar B VE. Shark cartilage-induced hepatitis. *Ann Intern Med* 1996;125(9)(Nov 1):780-1.
14. Miller DR, Anderson GT, Stark JJ, Granick JL, Richardson D. Phase I/II trial of the safety and efficacy of shark cartilage in the treatment of advanced cancer. *J Clin Oncol* 1998;16(11):3649-55.
15. Lubalin JS, Harris-Kojetin LD. What do consumers want and need to know in making health care choices? *Med Care Res Rev*. 1999;56 Suppl 1:67-102